Department of Physics and Astronomy, University of Canterbury,
Private Bag 4800, Christchurch, New Zealand

Averaging the Inhomogeneous Universe

YAO Hui

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

MAPH480 Project 2006

Supervisor: Dr David L. WILTSHIRE

Abstract

We re-formulate and examine T. Buchert's recent averaging scheme for scalars in cosmological applications of general relativity. The equation thus obtained can be used to describe the averaged quantities of an arbitrary inhomogeneous co-moving region and show the importance of back-reaction.

We also study the use of information theory in this averaging framework. Original extensions are mainly made along two lines: the information of inhomogeneity for different scales are compared; the possibility of use of Shannon's entropy in inhomogeneous cosmology are investigated. We also discuss the non-locality of gravitational energy in inhomogeneous cosmology.

Examples of cosmological solutions of Buchert's averaging scheme are studied.

Contents

Chapter 1

Introduction

The main direction of modern fundamental physics is to conquer the unknown in the two extreme scales. On the one hand, pursuits are being continued to understand the final building blocks of the Universe and the nature of spacetime at Planck scale. On the other, it is the modernisation of cosmology—the study of the Universe at the large scale—that made these two efforts symmetric and, more importantly, connected, as the two extreme ends will turn out to be the one: unified laws of physics, governing all.

This effort is a manifestation of belief in the simplicity of nature. However, caution has also been exerted. Einstein said: "Everything should be made as simple as possible, but not simpler.[1]" Following in this spirit, we study cosmology—the simple laws describing the most complex object; in particular, we study inhomogeneous cosmology, the class of theories which are far more complex than the simple homogeneous standard models; even more particularly we will study the averaging procedure in inhomogeneous cosmology, the powerful and profound technique which enables us to extract simple properties from a complex system and to examine the "Universe seen at different scales". We therefore, before our formal exposition, ask three questions: What is the current status of cosmology? Why inhomogeneous cosmology? Why averaging in inhomogeneous cosmology?

What is the current status of cosmology?

The advancement of modern cosmology is brought forth by technology and the availability of massive flow of astronomical data, especially those from Sloan Digital Sky Survey(SDSS), the Hubble telescope, the Chandra telescope and WMAP. Most cosmologists would agree on that we are entering an era of "precision cosmology", when data collected will be made more and more precise to put constraint on cosmological parameters, cosmological models and new physics proposed.

Although being in an optimistic situation, we need to emphasize that cosmology is a very special subject and therefore has its special difficulties [2]. Its object of study, the Universe as a whole together with its role as the background for all the rest of physics and science, challenges both physicists and philosophers; philosophical preferences strongly influence the resulting understanding and choices of models studied. Fashion or sociology in science could be a tremendous negative force in studies of theory and investments of experiments, and analysis of data is inevitably model-dependent. We are limited in our ability to observe both

[1] As quoted in [1].

the very distant regions and the very early times, and limited in our ability to test physics in a direct manner over the large scale and to test high energy physics relevant at the earliest epochs. It is very impressive to see how the new version of the standard model, the ΛCDM model, is able to explain and predict wide classes of independent phenomena, especially the newly released WMAP data [3]. However, "many commonly discussed elements of cosmology still are on dangerous ground" [4]. Here we discuss some conceptual problems [5] of modern cosmology which are of particular interest to our later discussion.

1) **Gravitational energy problem.** In FLRW model, conservation of energy-momentum in General Relativity(GR) is equivalent to $dE = -p\,dV$ for an arbitrary co-movig region. This of course should be directly interpreted as the first law of thermodynamics for fluid with no heat transfer. As long as the fluid is not ideal dust with zero pressure, we see that an expanding universe would lead to a decrease of matter energy. We will discuss these problems in section 3.4.

2) **Dark energy problem.** In the past decade, observations of the luminosity-distance of TypeIA supernovae have been interpreted within FLRW model as that the scale factor of the Universe must have a positive second order time derivative ("the Universe is accelerating"). Together with WMAP and other independent observations, the best-fit parameters indicate that 75% of the energy content of the Universe must be in a mysterious form of non-luminous, non-gravitationally-clumping, negative pressure fluid, known as "dark energy". The nature of dark energy so far still lacks physical basis.

3) **Hubble–de Vaucouleurs paradox.** Hubble's discovery that the redshift of a galaxy is proportional to its distance is one of the cornerstones of the standard model. According to modern observations, a linear Hubble's law is well established starting from scales about 1.5 to 2 Mpc. This is usually understood to imply a homogeneous Universe. Hubble's law is directly deducible from the FLRW model, but the converse is not true. In fact, an inhomogeneous matter distribution is observed from the scale of the atoms up to at least the scales of superclusters and voids. Sandage *et al.* [6] were the first to note the puzzling co-existence of the linear Hubble's law and the (at least local) inhomogeneity. Although their original paper was arguing against de Vaucouleurs' hierarchy Universe based on the linearity of Hubble's law, it remains a great challenge to explain the co-existence of these two independent universal aspects and it is indeed a crucial task to solve in modern cosmology.

Why inhomogeneous cosmology?

The central idea of inhomogeneous cosmology is to emphasize the importance of cosmological structures. There are three aspects of these:

- To study classes of (exact or non-exact) solutions of Einstein's equations as cosmological models without the assumptions of homogeneity and isotropy. These problems should not be categorised merely as pure mathematical problems. In fact, exact inhomogeneous solutions can describe several features of the Universe in agreement with observations, especially the existence of voids [6].

- To investigate the effects of global expansion on local physics.

- To study effects of local inhomogeneous structures on the dynamics of global geometry. Termed "back-reaction", this aspect of cosmology is attracting more and more researchers currently. This is encouraged by the increasing evidence of inhomogeneity and motivated as alternative solutions to the dark energy problem. We will be particularly interested in this last aspect.

One particular interesting idea is the modern version of the hierarchy Universe—fractal cosmology. Fractals are ubiquitous in Nature; should cosmology be an exception? Fractal cosmology is based on modern observation. Pietronero [8], using statistical methods, argues that in the various surveys galaxy counts are proportional to r^D at least up to 20Mpc, where r is radial distance and $D \approx 2$ is fractal dimension. Recently it was further argued that the above result is consistent with SDSS [9]. Fractal cosmology is based on a weaker interpretation of the Copernican principle—the "conditional cosmological principle" formulated by Mandelbrot. A concrete formulation of these ideas in GR is very difficult. Examples of these efforts are [10], [11]. The validity of fractal cosmology remains an interesting open question.

Why averaging in inhomogeneous cosmology?

Any mathematical description of a physical system depends on an averaging scale characterizing the nature of the model [12]. This scale is usually taken as understood and therefore hidden from the model, but it is one crucial element of the model. For example, a fluid continuum is an averaged concept over a scale which must be large enough such that the property of each individual molecule can be neglected, yet small enough such that spatial gradients of properties under interest are well represented and not smoothed out. Similar concerns were also pointed out by Tolman [13] back in the 1930s: although most cosmological quantities are assumed to be smooth functions which assign an exact value to each point of the spacetime, they are really macroscopically and phenomenologically identified. It is the (spacetime) averaged quantity that have direct observational status and physical meanings [14].

Applications of general relativity in cosmology usually start from implicitly assuming the validity of the theory at the largest scale, whereas GR is indeed only tested directly at the solar scale. The problem lies in the fact that GR is not scale invariant! Suppose we have some well defined averaging procedure $\langle\,\rangle_1$ over some scale L_1 for tensors and assume GR is valid in a direct manner over L_1, by which we mean the following: (i) calculate the Einstein's tensor $\mathbf{G}$ from the L_1-averaged metric tensor $\langle\mathbf{g}\rangle_1$ to obtain $\mathbf{G}(\langle\mathbf{g}\rangle_1)$; (ii) determine the L_1-averaged energy-momentum $\langle\mathbf{T}\rangle_1$; (iii) then Einstein's equations tell us that $\mathbf{G}(\langle\mathbf{g}\rangle_1) = \kappa\,\langle\mathbf{T}\rangle_1$, where κ is a universal constant. In practice, L_1 is implicitly assumed and $\langle\,\rangle_1$ is normally dropped. Einstein's equation is then commonly written as $\mathbf{G}(\mathbf{g}) = \kappa\,\mathbf{T}$. However, at some larger scale L_2, we need to average the above equation to obtain: $\langle\mathbf{G}(\langle\mathbf{g}\rangle_1)\rangle_2 = \kappa\,\langle\langle\mathbf{T}\rangle_1\rangle_2$. If we pay attention to the fact that $\mathbf{G}$ is a function of $\langle\mathbf{g}\rangle_1$ and its first/second order derivatives, we see that the above valid equation is very different from $\mathbf{G}(\langle\langle\mathbf{g}\rangle_1\rangle_2) = \kappa\,\langle\langle\mathbf{T}\rangle_1\rangle_2$. The use of Einstein's equation implicitly requires an averaging scale L_1; its direct use over some scale L_1 is not in general mathematically compatible with its direct use over some other scale L_2.

As pointed out by Zalaletinov [15], we can either interpret GR macroscopically and derive GR itself from some microscopic equations describing systems at smaller scales, or we can interpret GR microscopically and derive the averaged macroscopic equations describing physical systems at larger scales. The latter interpretation is inevitable in the study of cosmology, since whatever scale we start with, we will need the averaged equations to describe the largest physical object in the Universe—the Universe itself. The purpose of such an averaging scheme in inhomogeneous cosmology is to obtain a best–fit smooth flow from an intrinsically lumpy physical system, so that the global average parameter can be compared with observation, as illustrated [16] schematically in Fig. 1.1:

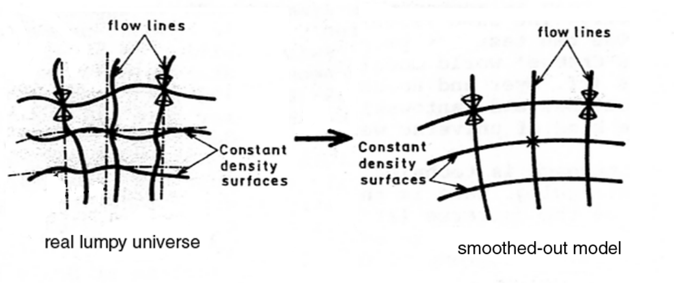

Figure 1.1: *Seeking a best–fit smoothed–out model*

By performing such an averaging procedure, irrelevant details of matter and geometry fluctuations are smoothed out, so that we can study the large scale qualitative behaviors of the system, as illustrated [16] schematically below:

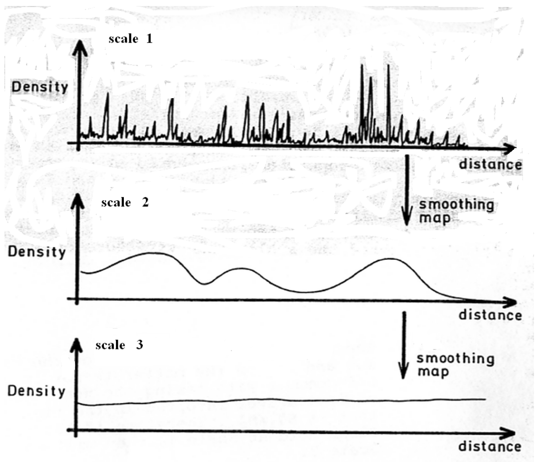

Figure 1.2: *A comparison of the same region at different scales.*

One simple approach of deriving the averaged macroscopic equations from an microscopic interpretation of GR is recently proposed by T. Buchert. This project is mainly based of a review of Buchert's recent work on averaging inhomogeneous cosmology, especially that of [17], [18], [19]. The formulation has been organized in a concise and coherent way and the calculations and assumptions have been examined. More importantly, comments and comparisons with other work are provided whenever possible, as encouragement for critical thinking in further studies. My main original contributions are made in section 3.2 and section 3.3, where the possibility of extending further use of information theory in inhomogeneous cosmology are examined. My understanding of the non-locality of gravitational energy are discussed in section 3.4.

Chapter 2

The Averaging Procedure

In this chapter we will study the averaging procedure recently proposed by Buchert [17]. Through out this paper, we will restrict ourselves with irrotational dust continuum, that is, cosmic fluid with negligible pressure and vorticity. For a similar averaging method covering general perfect fluid, see [20]. For an alternative averaging procedure, see [14]. For a general review on the early work of averaging problem prior to 1997, see chapter 8 of [7].

2.1 The 3+1 decomposition and Einstein's equations

We start with the 3+1 splitting of the 4-dimensional manifold of spacetime of the Universe $\mathcal{M}$ into one-parameter foliations of spacelike hypersurfaces Σ_t, labelled by global time coordinate[1] t. This is possible in general if $\mathcal{M}$ is globally hyperbolic. For such a choice of foliation, it is always possible to choose *Gaussian normal coordinates*[2], i.e. $ds^2 = -dt^2 + g_{ij}dx^i dx^j$. The parameter t is the proper time of observers co-moving with the cosmic fluid, for whom $\partial_t x^i = 0$. The trajectories of fluid elements follow timelike geodesics everywhere orthogonal to Σ_t, with unit tangent 4-vector $u^a = (1,0,0,0)$, $u_a = (-1,0,0,0)$.

For this choice of foliation, we define the *extrinsic curvature* of Σ_t with respect to $\mathcal{M}$:

$$K_{ij} \equiv -u_{i;j} = \Gamma^0{}_{ij}\, u_0 = -\frac{1}{2} g_{ij,0} \tag{2.1}$$

For this specific choice of coordinates, Einstein's equations with a dust source and a cosmological constant $G^{ab} = 8\pi\rho u^a u^b - \Lambda g^{ab}$ are equivalent to (see, e.g., section II of [23]):

$$\frac{1}{2}(R + K^2 - K^i{}_j K^j{}_i) = 8\pi G\rho + \Lambda \tag{2.2}$$

$$K^i{}_{j||i} - K_{,j} = 0 \tag{2.3}$$

$$K^i{}_{j,0} = K K^i{}_j + R^i{}_j - (4\pi G\rho + \Lambda)\delta^i{}_j \tag{2.4}$$

where $R = R^i{}_i$ is the Ricci scalar corresponding to the metric g_{ij}, $K = K^i{}_i$, and indices are lowered/raised with respect to g_{ij}.

[1] See section 2.1 of [21] and reference therein, or section 21.4, 21.5 of [22] for further discussion.

[2] This will be our working assumption, which is also the assumption of FLRW model and various other models. But it introduces certain physical and geometrical restrictions, especially the existence of a global proper time, which should be investigated more carefully.

For irrotational dust, we decompose the extrinsic curvature as: $-K_{ij} = u_{i;j} \equiv \sigma_{ij} + \frac{1}{3}\theta g_{ij}$, where $\theta \equiv u^i{}_{;i}$ is the *expansion rate*, and σ_{ij} is the *shear tensor* which is easily seen to be trace-free and symmetric. We define the *rate of shear* σ correspondingly as $\frac{1}{2}\sigma^i{}_j\sigma^j{}_i$. (See, e.g., chapter 22 of [22]. Note the *four-acceleration*: $u^a u^b{}_{;a}$ vanishes.)

With the above notation, (2.1),(2.2),(2.3) now read

$$g_{ij,0} = 2\sigma_{ij} + \frac{2}{3}\theta g_{ij} \tag{2.5}$$

$$\frac{1}{2}R + \frac{1}{3}\theta^2 - \sigma^2 = 8\pi G\rho + \Lambda \tag{2.6}$$

$$\sigma^i{}_{j||i} - \frac{2}{3}\theta_{,j} = 0 \tag{2.7}$$

where all quantities depend on 4-coordinates (x^a). We will call (2.6) the *Friedmann equation*. Contracting i, j of (2.4), with the help of (2.6), we have the following *Raychaudhuri equation*:

$$\dot{\theta} + \frac{1}{3}\theta^2 + 2\sigma^2 + 4\pi G\rho - \Lambda = 0 \tag{2.8}$$

Using again (2.6) and (2.8), equation (2.4) now reads

$$\sigma^i{}_{j,0} + \theta\sigma^i{}_j + R^i{}_j - \frac{1}{3}R\delta^i{}_j = 0 \tag{2.9}$$

Finally, the conservation of energy equation $(\rho u^a u^b)_{;a} = 0$, when contracted with u_b, gives:

$$\dot{\rho} = K\rho = -\theta\rho \tag{2.10}$$

We also denote $J = \sqrt{det(g_{ij})}$, then we have

$$\dot{J} = \frac{1}{2}g^{ik}g_{ki,0}J = \theta J \tag{2.11}$$

It then follows from (2.10) and (2.11)that:

$$\rho(t, x^i) = \rho(t_0, x^i)J(t_0, x^i)(J(t, x^i))^{-1} \tag{2.12}$$

2.2 Averaging Einstein's equations

We define the spatial averaging of a scalar field $\Psi(t, x^i)$ over a compact portion D of spacelike hypersurface Σ_t as the following linear operation:

$$\Psi(t, x^i) \to \left\langle \Psi(t, x^i) \right\rangle_D \equiv \frac{1}{V_D}\int_D \Psi(t, x^i)dV \tag{2.13}$$

where we have used $dV \equiv Jd^3x$ for clarity and $V_D \equiv \int_D dV$ is the integrated proper volume.

We therefore obtain the averaged covariant (with respect to spatial coordinates transformation) scalar field if we assign each point of D with the value $\langle\Psi(t, x^i)\rangle_D$ indistinguishably. However, such obtained "coarse-grained" field would not be globally smooth, and would depend on arbitrary choice of division of Σ_t into a collection of averaging regions. Alternatively, we can proceed as following: for some t, at each point $x^i \in \Sigma_t$, we assign an unique co-moving averaging region D_{x^i} with x^i in its interior. We define the value of the averaged field at x^i to be:

$$\langle\Psi\rangle(x^i,t) \equiv \frac{1}{V_{x^i}} \int_{D_{x^i}} \Psi(t,x^i)\, dV$$

where again V_{x^i} is the proper volume of the co-moving region at time t. In order to get an unique, covariant, globally smooth field, we would have to require a way of "coordinating" averaging regions. For further pursuit along this direction, see cite15. However, from now on, unless otherwise clearly stated, we will constrain ourselves with the study of a given fixed co-moving averaging region. We emphasise that the averaged quantity is domain dependent, even if when subscript "D" is dropped for convenience.

We also remark that the averaging defined above has physical meaning most transparent if our primary concern of Ψ is energy density ρ. But for an arbitrary scalar field, the physical meaning of this procedure is not as clear. For example, for a scalar field of temperature, it is not clear that the volume averaging represents "the averaged temperature". For energy density field, the local form of conservation of energy (2.12) is equivalent to

$$M \equiv \int \rho\, dV = \langle\rho\rangle V = const. \tag{2.14}$$

However, we make the warning that the total matter energy for an isolated co-moving region is not a constant in general. See section 3.4 for further discussion.

In analogy to the standard cosmology, we define an *effective scale factor*

$$a(t) = (\frac{V(t)}{V_0})^{\frac{1}{3}} \tag{2.15}$$

where a subscript "0" denotes the value at some initial time t_0.

We similarly define *effective Hubble function*:

$$H(t) = \frac{\dot{a}}{a} \tag{2.16}$$

and *effective deceleration parameter*:

$$q(t) = -\frac{\ddot{a}}{a}\frac{1}{H^2} \tag{2.17}$$

These quantities of course reduce back to the standard definition when we have, hypothetically, a homogeneous energy distribution.

With the above notation, the averaged expansion rate has a transparent meaning:

$$\langle\theta\rangle = \frac{\dot{V}}{V} = 3\,\frac{\dot{a}}{a} = 3H \tag{2.18}$$

The important *commutation rule* then follows naturally from (2.11),(2.18):

$$\partial_t\langle\Psi\rangle - \langle\partial_t\Psi\rangle = \langle\Psi\,\theta\rangle - \langle\Psi\rangle\langle\theta\rangle = \langle\Psi\,\delta\theta\rangle = \langle\theta\,\delta\Psi\rangle = \langle\delta\Psi\,\delta\theta\rangle \tag{2.19}$$

where $\delta\Psi = \Psi - \langle\Psi\rangle$ and $\delta\theta = \theta - \langle\theta\rangle$. The key statement of (2.19) is that the operations of spatial averaging and time evolution do not commute.

For $\Psi = \rho$, this gives:

$$\partial_t \langle \rho \rangle = - \langle \rho \rangle \langle \theta \rangle \tag{2.20}$$

which should be compared with (2.10).

We next obtain two important results. The *averaged Friedmann equation* is the spatial averaging of (2.6):

$$3\,(\frac{\dot{a}}{a})^2 - 8\pi G \frac{M}{V_0 a^3} + \frac{1}{2}\langle R \rangle - \Lambda = -\frac{\mathcal{Q}}{2} \tag{2.21}$$

where

$$\mathcal{Q} = \frac{2}{3}(\langle \theta^2 \rangle - \langle \theta \rangle^2) - 2\left\langle \sigma^2 \right\rangle = \frac{2}{3}\left\langle (\theta - \langle \theta \rangle)^2 \right\rangle - 2\left\langle \sigma^2 \right\rangle \tag{2.22}$$

is the *back-reaction term* which represents the global effect of local inhomogeneity. Averaging (2.8), using (2.18) and (2.19), we obtain the *averaged Raychaudhuri equation*:

$$3\,\frac{\ddot{a}}{a} + 4\pi G \frac{M}{V_0 a^3} - \Lambda = \mathcal{Q} \tag{2.23}$$

Equations (2.21) and (2.23) as well as their special forms in particular cosmology of chapter 4 will be referred simply as the *averaged equations* for convenience.

If we treat M, V_0 and Λ as parameters to be determined observationally, then (2.21) and (2.23) form a system of two ordinary differential equations with three unknowns. Unless we introduce extra simplifying assumptions, this system can not be solved.

We end this section by stating the following necessary *integrability condition*, obtained by differentiating (2.21) and using (2.23):

$$\partial_t \mathcal{Q} + 6\,\frac{\dot{a}}{a}\mathcal{Q} + \partial_t \langle R \rangle + 2\,\frac{\dot{a}}{a}\langle R \rangle = 0 \tag{2.24}$$

or equivalently:

$$\frac{1}{a^6}\,\partial_t(\mathcal{Q} a^6) + \frac{1}{a^2}\partial_t(\langle R \rangle\, a^2) = 0 \tag{2.25}$$

which shows that the averaged intrinsic curvature $\langle R \rangle$ and the averaged extrinsic curvature (encoded in $\mathcal{Q}$) are dynamically coupled.

2.3 The cosmic quartet

In this section, we draw further analogies to homogeneous cosmology and study the features of averaged equations for a fixed co-moving region.

We start by interpreting $\mathcal{Q}, \langle R \rangle, \Lambda$ as effective sources of gravitation, and define the corresponding *effective energy density* as:

$$\rho_{\mathcal{Q}} = -\frac{1}{16\pi G}\mathcal{Q} \qquad \rho_{\langle R \rangle} = -\frac{1}{16\pi G}\langle R \rangle \qquad \rho_\Lambda = \frac{1}{8\pi G}\Lambda \tag{2.26}$$

We also define the *total energy density* and the *critical density* as:

$$\rho_{tot} = \langle \rho \rangle + \rho_{\mathcal{Q}} + \rho_{\langle R \rangle} + \rho_\Lambda, \qquad \rho_{cri} = \frac{3H^2}{8\pi G} \tag{2.27}$$

We emphasize that although formally identical, the critical density defined here in general does not play the same role as the critical density in the standard model, i.e., it is not the borderline between a closed Universe and an open Universe.

We continue to define the corresponding *effective pressure* respectively as:

$$p_{\mathcal{Q}} = \rho_{\mathcal{Q}} = -\frac{1}{16\pi G}\mathcal{Q} \qquad p_{\langle R\rangle} = -\frac{1}{3}\rho_{\langle R\rangle} = \frac{1}{48\pi G}\langle R\rangle \qquad p_{\Lambda} = -\rho_{\Lambda} = -\frac{1}{8\pi G}\Lambda \tag{2.28}$$

$$p_{tot} = p_{\mathcal{Q}} + p_{\langle R\rangle} + p_{\Lambda} = -\frac{1}{16\pi G}\mathcal{Q} + \frac{1}{48\pi G}\langle R\rangle - \frac{1}{8\pi G}\Lambda \tag{2.29}$$

The significance of the above formalism lies almost entirely in the following simple and physically clear re-expression of (2.21),(2.23),(2.24):

$$3\,(\frac{\dot{a}}{a})^2 = 8\pi G\,\rho_{tot} \tag{2.30}$$

$$3\,\frac{\ddot{a}}{a} = -4\pi G\,(\rho_{tot} + 3\,p_{tot}) \tag{2.31}$$

$$\dot{\rho}_{tot} + 3H(\rho_{tot} + p_{tot}) = 0 \tag{2.32}$$

Equations (2.30),(2.31) of course are just normal Friedmann equation and Raychaudhuri equation with ordinary homogeneous fluid density replaced by our effective density which takes into consideration of inhomogeneity structure.

However, we remark that this interpretation has only limited use. Especially, $\rho_{tot}V$ for an isolated fixed co-moving region generally is not a constant with respect to time. This is manifest from differentiating $\rho_{tot}V$ and comparing to (2.25). There is no physical reason to view the change of $\rho_{tot}V$ as transference to gravitational energy. In particular we note that the term $\rho_{\langle R\rangle}$ and $\rho_{\mathcal{Q}}$ already represent gravitational structure. See section 3.4 for further discussion.

For completeness of the analogy, we now further define the corresponding *effective density parameter* respectively as following:

$$\Omega_m = \frac{\langle\rho\rangle}{\rho_{cri}} = \frac{8\pi GM}{3V_0 a^3 H^2} \qquad \Omega_{\mathcal{Q}} = \frac{\rho_{\mathcal{Q}}}{\rho_{cri}} = -\frac{\mathcal{Q}}{6H^2}$$
$$\Omega_{\langle R\rangle} = \frac{\rho_{\langle R\rangle}}{\rho_{cri}} = -\frac{\langle R\rangle}{6H^2} \qquad \Omega_{\Lambda} = \frac{\rho_{\Lambda}}{\rho_{cri}} = \frac{\Lambda}{3H^2} \tag{2.33}$$

and therefore (2.21) is equivalent to a form Buchert [18] dubs the *cosmic quartet*:

$$\Omega_m + \Omega_{\mathcal{Q}} + \Omega_{\langle R\rangle} + \Omega_{\Lambda} = 1 \tag{2.34}$$

Similar to the remark following (2.23), the system of ordinary differential equations (2.30),(2.31) could be solved if a *cosmic equation of state* is given in the form $p_{tot} = p_{tot}(\rho_{tot}, a)$. Therefore questions related to the evolution of the inhomogeneous Universe could be "reduced" to the problem of finding a cosmic state on a given spatial scale. Although formally similar to the situation in the standard cosmology, here the equation of state is dynamical and depends on the details of the evolution.

Chapter 3

The Information of Cosmological Inhomogeneity

Before we move on to investigate the interesting examples of the solutions to the system of equations obtained in chapter 2, in this chapter we will use the language of information theory to study the spatially averaged inhomogeneous Universe. In section 3.1 we review the work of Hosoya *et al.* [18]. In section 3.2 we point out one important feature of the quantity proposed in [18]. In section 3.3, we suggest another possible definition of [18].

3.1 The relative entropy in cosmology

In standard information theory, the well known *relative entropy* or *Kullback-Leibler information* S is defined [1] as

$$
\begin{aligned}
S(p||q) &= \sum p_i \, \ln(\frac{p_i}{q_i}) \qquad \text{for discrete random variable,} \qquad \text{or} \\
S(p||q) &= \int p(x) \, \ln(\frac{p(x)}{q(x)}) \, dx \qquad \text{for continuous random variable}
\end{aligned}
\tag{3.1}
$$

where p, q are two different probability distributions of the same random variable. This is a quantity measuring the inefficiency of assuming the distribution is q when the "true" distribution is p, or a measure of the "distance" between two different distributions. (The word "distance" is used in a very limited sense, since this quantity is not generally symmetric in p, q, let alone satisfying the triangle inequality.)

Application of this concept in inhomogeneous cosmology starts from identifying the probability distribution of p and q as energy density distribution of ρ and $\langle\rho\rangle$ respectively:

$$
S_D = \int_D \rho \, \ln(\frac{\rho}{\langle\rho\rangle_D}) \, dV \tag{3.2}
$$

We have dropped the explicit distribution dependence of S_D since they are implicitly assumed to be ρ, $\langle\rho\rangle_D$ without chances of confusion, but we have put subscript D to emphasize its regional dependence, for reasons to become clearer in the next section. This quantity is strictly positive, unless in the case of a homogeneous distribution $\rho = \langle\rho\rangle_D$ it vanishes. Similar to the original use of relative entropy in information theory, we could view (3.2) as an estimate of the inefficiency of assuming a homogeneous region when the true distribution is inhomogeneous, or the degree of cosmological inhomogeneity.

Comparing with the standard index of inhomogeneity in cosmology:

$$\frac{(\Delta\rho)^2}{\langle\rho\rangle_D} = \frac{\langle\rho^2\rangle - \langle\rho\rangle^2}{\langle\rho\rangle_D} \tag{3.3}$$

we see the relative entropy is at least complementary. As pointed out in [18], (3.2) and (3.3) are both members of a one-parameter family of inhomogeneity measures, the *Tsallis relative entropy*, defined by :

$$\mathcal{F}_\alpha = \frac{\langle\rho\rangle_D}{\alpha}\left(\left\langle(\frac{\rho}{\langle\rho\rangle_D})^{\alpha+1}\right\rangle_D - 1\right) \tag{3.4}$$

with α as a real parameter. When $\alpha = 1$, (3.4) reduces to (3.3) and when $\alpha \to 0$, the limit is (3.2). However, the relative entropy we adopted as in (3.2) is the only quantity satisfies the following remarkable property: differentiating (3.2) and use (2.19), we have:

$$\frac{\dot{S}_D}{V_D} = \langle\partial_t\rho\rangle_D - \partial_t\langle\rho\rangle_D = -\langle\rho\delta\theta\rangle_D = -\langle\theta\delta\rho\rangle_D = -\langle\delta\rho\delta\theta\rangle_D \tag{3.5}$$

We interpret this as: the production rate of relative information per unit volume is the source of non-commutativity of spatial averaging and time evolution. This justifies the use of (3.2) as an important quantification of inhomogeneity. Using Schwarz inequality:

$$\left(\int f(x)\,g(x)\,dx\right)^2 \leq \int f(x)^2\,dx \int g(x)^2\,dx \tag{3.6}$$

we have:

$$\left|\frac{\dot{S}_D}{V_D}\right| = |\langle\delta\rho\,\delta\theta\rangle_D| \leq \Delta\rho\Delta\theta \tag{3.7}$$

where

$$\Delta\rho = \sqrt{\langle(\delta\rho)^2\rangle_D} \qquad \text{and} \qquad \Delta\theta = \sqrt{\langle(\delta\theta)^2\rangle_D} \tag{3.8}$$

That is, the production rate of relative information per unit volume is bounded by the product of amplitudes of density and expansion fluctuation. Hosoya *et al.* interpret this as a competition between the production of information in the Universe and its volume expansion.

We now follow [18] to study the condition upon which the second time derivative of (3.2) is positive, the *time convexity* of relative entropy. Differentiating (3.5) as well as using (2.19), we have:

$$\frac{\ddot{S}_D}{V_D} = -\left\langle\delta\rho\,\delta(\dot{\theta})\right\rangle_D + \langle\rho\rangle_D(\Delta\theta)^2 \tag{3.9}$$

Recalling (2.8) and using the variation of Schwarz inequality, $\langle\delta a\,\delta b\rangle_D \geq -\Delta a\Delta b$, we have:

$$\begin{aligned}\frac{\ddot{S}_D}{V_D} &= 4\pi G(\Delta\rho)^2 + \langle\rho\rangle_D(\Delta\theta)^2 + \frac{1}{3}\left\langle\delta\rho\,\delta(\theta^2)\right\rangle_D + 2\left\langle\delta\rho\,\delta(\sigma^2)\right\rangle_D \\ &\geq 4\pi G(\Delta\rho)^2 + \langle\rho\rangle_D(\Delta\theta)^2 - \Delta\rho(\frac{1}{3}\Delta(\theta^2) + 2\Delta(\sigma^2))\end{aligned} \tag{3.10}$$

Viewing the right hand side as quadratic in $\Delta\rho$, we obtain the sufficient condition for $\ddot{S}_D$ to be strictly positive:

$$\sqrt{4\pi G \langle\rho\rangle_D} > \frac{\frac{1}{3}\Delta(\theta^2) + 2\Delta(\sigma^2)}{2\Delta\theta} \tag{3.11}$$

That is, time convexity of relative entropy will be obtained if gravity dominates over expansion and shear fluctuations. Time convexity implies that the rate of structure formation eventually increases.

Looking back at (3.5), we see that relative entropy would be produced if, on average, overdense fluid element ($\delta\rho > 0$) are contracting ($\delta\theta < 0$), or underdense element ($\delta\rho < 0$) are expanding ($\delta\theta > 0$). In structure formation, the processes of relative accumulation of matter (cluster formation) and relative dilution of matter (void formation) create an asymmetry of states, for large enough time, at certain scales. Primarily motivated by this, in [18], Hosoya *et al.* proposed the following *Conjecture*: The global relative entropy[1]of a dust model S_{Σ_t} is an increasing function of time, for sufficiently large time. At least at the time of this report is written, the proof of this conjecture is not completed.

It is contemplated in [18] that the relative entropy proposed above may turn out to play a fundamental role also in many other aspects of gravity, e.g., for the study of black holes and the early Universe. Such possibly deep implications may remain future work, however, in the next two sections, we restrict ourselves to two possible modest extensions of their work.

3.2 The component information and the structure information

The use of the word "measure" in the literature may sometimes cause confusion. In [18] it is implied that the relative entropy is a *measure* in the sense standard measure theory and probability, which is defined [24] as a function μ from a σ-algebra over a set $\mathcal{M}$ into[2] $[0, \infty]$ such that: if $\{D_i\}$ is a countable sequence of pairwise disjoint sets in $\mathcal{M}$, then $\mu(\bigcup_{i=0}^{\infty} D_i) = \sum_{i=0}^{\infty} \mu(D_i)$. Contrary to [18], we point out that S in (3.2) is not a *measure* in the above sense, when viewed as a function of different averaging regions. In this section and the next, we temporary drop the restriction of averaging over a fixed domain and study some properties of averaging over different regions and scales.

The essential property of a measure is that an object can be measured by breaking it up to smaller pieces, measuring those, and adding the result. The above defined function is not additive, in fact:

$$S - \sum_{i=0}^{n} S_i = \sum_{i=0}^{n} V_i \langle\rho\rangle_i \ln(\frac{\langle\rho\rangle_i}{\langle\rho\rangle}) \tag{3.12}$$

where a subscript i represents the corresponding quantity of region D_i, and the quantity

[1]By "global" we mean the relative entropy coresponds to the whole compact spacelike hypersurface Σ_t. This assumption of compactness of Σ_t, due to Hosoya *et al.* , of course needs to be justified.

[2]∞ is a symbol such that $a < \infty,\ a + \infty = \infty + a = \infty,\ \forall a \in \mathbb{R}$.

without subscript is that of region $\bigcup_{i=0}^{n} D_i$. Also, as before:

$$\langle\rho\rangle = \frac{1}{V}\int \rho dV = \frac{\sum_{i=0}^{n} V_i \langle\rho\rangle_i}{\sum_{i=0}^{n} V_i} \tag{3.13}$$

The equation (3.13) is non-negative; it vanishes iff $\langle\rho\rangle_i = \langle\rho\rangle$, for $\forall i$. This could be generalized to the case when $n \to \infty$, either countably or not, which shows that S in (3.2) is not additive in general.

It is not the subtlety of the usage of mathematical terminology that is our main concern. Comparing the remarkable similarity between (3.2) and r.h.s. of (3.12), the physical meaning of this is transparent: the total information of a system (S, the first term of l.h.s. of (3.12)) is different from the sum of the *component information* (S_i in the second term of l.h.s. of (3.12)), with the difference being the *structure information* (the r.h.s of (3.12)) of its broken components D_i. This is manifest even for a simple system when the only property under interest is gravity. (By analogy, breaking a picture would leads to a greater loss of information.)

As argued in [25] and [12], the loss of information is closely related to the coarse-grained multi-scale description of matter and gravity. If we divide the Universe (or at least a compact proportion of it) as the union of regions of the same scale, with the total information decomposed as the sum of component information and structure information, then the finer the description is, the more structure information we would have, and correspondingly the less component information would remain. In analogy to that the individual pixels of a picture convey no information, we also note that the component information is actually hidden behind the averaged density; it is the (spacetime) averaged quantity that have direct observational status and physical meanings [14]. We illustrate these ideas by the following schematic diagram:

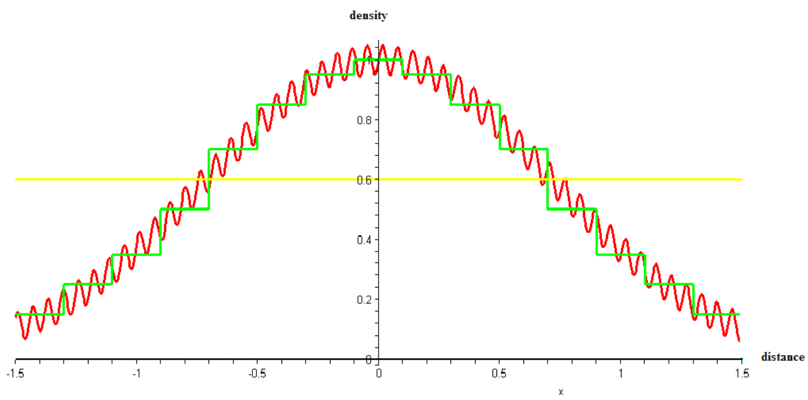

Figure 3.1: *Energy density of a region averaged over different scales, where only one of the three spatial dimensions are plotted. The red curve represents a detailed density distribution with its information of inhomogeneity given by (3.2). The green curve represents a density distribution averaged over a intermediate scale with only the structure information left. The yellow curve represents the density distribution averaged over the whole region with zero information of inhomogeneity.*

Before we finish this section, we note that, although being not a mathematical concept of a measure, S in (3.2) still possess the following physically desirable properties:

- $\lim_{V_D \to 0} S_D(\rho) = 0$. Thus we can use the notation $S_\emptyset = 0$, where $\emptyset$ is an empty subset of $\mathcal{M}$.

- $S_D \leq S_{D'}$, if $D \subseteq D'$.

Finally, we show for two arbitrary sets A and B, possibly non-disjoint, the following relation holds:

$$S_{A\cup B} - S_A - S_B = V_A \langle\rho\rangle_A \ln(\frac{\langle\rho\rangle_A}{\langle\rho\rangle_{A\cup B}}) + V_B \langle\rho\rangle_B \ln(\frac{\langle\rho\rangle_B}{\langle\rho\rangle_{A\cup B}}) - \int_{A\cap B} \rho \, \ln(\frac{\rho}{\langle\rho\rangle_{A\cup B}}) \, dV \quad (3.14)$$

where when $A \cap B = \emptyset$, this reduces to (3.12) with $n = 2$.

3.3 The effective entropy in cosmology

In [18], it is claimed the expression (3.2) was deduced. What seems to be the real case is that Hosoya *et al.* first applied the concept of *relative entropy* from standard information theory [1] in a cosmological context and then justified the physical importance of this definition by the remarkable relation of (3.5). As pointed out in section 3.1, application of this concept in cosmology starts from the identification of the probability distribution with matter density distribution. Following the same ideas, it's natural to apply also the concept of *Shannon entropy*[3] from information theory:

$$\begin{aligned} \mathbb{S}(p) &= -\sum p_i \, \ln(p_i) \qquad && \text{for a discrete variable,} \qquad \text{or} \\ \mathbb{S}(p) &= -\int p(x) \, \ln(p(x)) \, dx \qquad && \text{for a continuous variable} \end{aligned} \quad (3.15)$$

in a similar way by defining the *effective entropy*, or simply *entropy* for short, as[4]:

$$\mathbb{S}_D(\rho) = -\int_D \rho(x) \, \ln(\frac{\rho(x)}{\rho_0}) \, dV \quad (3.16)$$

where ρ_0 is a constant with the dimension $Joule/Volume$, as opposed to the region–dependent and time–dependent $\langle\rho\rangle$ in [18]. This needs to be determined from appropriate physics. If we normalise ρ_0 to be $\langle\rho\,\rangle_{\Sigma_{t_0}}$ for a compact global hypersurface Σ_{t_0} at some fixed initial time t_0, then $-\mathbb{S}_{\Sigma_{t_0}}(\rho) = S_{\Sigma_{t_0}}$. However, as we now demonstrate, Shannon's entropy thus interpreted in cosmology possesses interesting physical properties independent of the choice of ρ_0.

First of all, we notice that $\mathbb{S}_D(\rho)$ is not positive-definite, which is in accord with the standard information theory for continuous variable. It is not the second law of thermal dynamics that requires a positive entropy; it is convention and the third law of thermal

[3] Strictly speaking, Shannon entropy only refers to the case of a discrete variable. The generalization of Shannon entropy for a continuous variable is called *differential entropy* or sometimes *Boltzmann entropy*. There are important differences between the two, however, they are not our concern here. See, e.g., chapter 9 of [1].

[4] For the physical dimension to be the same as that of ordinary matter entropy, we should have defined $\mathbb{S}_D(\rho) = -const \int_D \rho(x) \, \ln(\frac{\rho(x)}{\rho_0}) \, dV$. Here *const* is a constant with the dimension of $Kelvin^{-1}$, e.g. $k\sqrt{\frac{G}{\hbar c^5}}$. This remark also applies to the relative entropy S defined in [18].

dynamics that gives zero entropy for zero absolute temperature state of matter and positive otherwise [13]. We see no reason to carry on this convention to the effective entropy of cosmological structure. In fact, we have the following limit:

$$\lim_{\rho \to 0} \mathbb{S}_D(\rho) = 0 \tag{3.17}$$

We also notice that $\mathbb{S}_D(\rho)$ is additive, i.e.;

$$\mathbb{S}_{A \cup B}(\rho) = \mathbb{S}_A(\rho) + \mathbb{S}_B(\rho), \quad \text{for all } A \cap B = \emptyset \tag{3.18}$$

This is totally as expected. The physical reason for the difference between here and (3.12) is as following: the entropy in (3.16) for the exact density distribution ρ is a measure of internal physical properties which only depends on ρ, we therefore expect for two isolated system without physical interactions, the total entropy is simply the sum of the entropy of the two sub-regions; on the other hand, the relative entropy in (3.2) crucially depends on the averaging region as a whole.

By comparison, for a system with two distinct components of matter: $\rho = \rho_1 + \rho_2$, we have

$$\mathbb{S}_D(\rho) > \ \mathbb{S}_D(\rho_1) + \mathbb{S}_D(\rho_2) \tag{3.19}$$

Comparing with (3.18), we see although the effective entropy is additive with respect to its regional dependence, but it is not additive with respective to its constitute dependence. Evidently the mixture of two components would lead to an increase of the total entropy.

We next calculate:

$$\mathbb{S}_D(\langle \rho \rangle_D) - \mathbb{S}_D(\rho) = S_D \tag{3.20}$$

i.e., the difference of the entropy corresponding to averaged matter distribution and "true" matter distribution is the relative entropy of the region. Notice that this result is not generally true for two arbitrary probability distributions in an information theory context. The consequence of (3.20) is that for a co-moving region at a certain time, the maximum entropy corresponds to a homogeneous matter distribution of the region. The coarse-grained version of (3.20) reads:

$$\mathbb{S}(\langle \rho \rangle) - \sum \mathbb{S}_i(\langle \rho \rangle_i) = \sum V_i \langle \rho \rangle_i \ln(\frac{\langle \rho \rangle_i}{\langle \rho \rangle}) \tag{3.21}$$

which, again, indicates that the coarser the description, the higher the entropy. This expression of course reduces back to (3.20) in the case $V_i \to 0$.

Using (2.10),(2.11),(2.18) and (2.19), we find that:

$$\dot{\mathbb{S}}_D = V_D \langle \rho \, \theta \rangle_D \tag{3.22}$$

$$\ddot{\mathbb{S}}_D = V_D \left\langle \rho \, \dot{\theta} \right\rangle_D \tag{3.23}$$

The expression (3.22) is positive for an expanding region with $\langle \rho \, \theta \rangle > 0$, regardless of its inhomogeneity. This regional dependence of the sign of $\dot{\mathbb{S}}$ is a key feature of an open system, where gravity is long range. If the whole spacelike hypersurface Σ_t of the Universe is compact, we would require $\langle \rho\, \theta \rangle_{\Sigma_t} > 0$ to satisfy the second law of thermodynamics for the effective entropy; otherwise caution is in order with any statements.

It is important to distinguish between two cases. At a local structure formation scale, it has been recently shown [26] that, although the inhomogeneity of density and temperature due to gravitational contraction results in a decrease of entropy, this is outweighed by an increase of entropy due to the increase of thermal energy from contraction. Hence the total thermal entropy satisfy the second law of thermodynamics without the necessity of introducing a gravitational entropy. On the other hand, in our case, local ordinary thermal entropy has been neglected, due to the choice of co-moving coordinates, which is a suitable approximation for ideal irrotational dust continuum. Our primary concern is the effective entropy at a cosmological scale.

Finally, for small density fluctuations, we have:

$$\frac{\dot{\mathbb{S}}_D}{V_D \, \langle \rho \rangle_D} \sim \langle \theta \rangle_D \,, \qquad \frac{\ddot{\mathbb{S}}_D}{V_D \, \langle \rho \rangle_D} \sim \left\langle \dot{\theta} \right\rangle_D \tag{3.24}$$

We end this chapter by considering some deep issues we have not resolved. The Decomposition of relative entropy into the sum of component information and structure information as in section 3.2 is of physical significance. The effective entropy proposed in section 3.3 is a natural extension of [18] and exhibits many physically desirable properties. However, numerous problems remain: How is the effective entropy related to the entropy of gravitational field? How is it related to the first law of relativistic thermodynamics, especially that formulated by A.Einstein, R.Tolman [13] which take into consideration of both ordinary fluid energy and gravitational energy. How can the concept of effective entropy be used to explain structure formation and the attractive nature of gravity? The increase of entropy in (3.21) relies on the expansion, and a homogeneous distribution actually corresponds to a higher entropy states as in (3.19); is this really the case? Is the analogy between the inequality (3.7) and the uncertainty relationship in quantum mechanics purely formal and accidental or could it be used as a hint of inspiration showing that there is some kind of "dual" relationship between energy density and expansion rate? How can we extend our theory to other aspects of physics, as contemplated by Hosoya *et al.* in [18], e.g., in the study of the early Universe when the content of fluid could not be modeled by dust, in the study of black holes and in the profound question of the arrow of time?

The important aspect of introducing a new language is certainly not to play with formalism, but to examine the possibility of solving problems remained from old theory, or the possibility of predicting new phenomena. But we also add that introducing a new language may lead to new insights and understanding of new principles. Without being patient with the infancy of new ideas, we could seldom advance much. Despite the numerous problems remained, we conclude the direction pointed out in this chapter worth further study.

3.4 Non-locality of Gravitational Energy

Understanding gravitational energy is of fundamental importance in gravitational physics and it is intimately related to many aspects of cosmology. In this section, we will discuss some aspects of the issue. We will still restrict ourselves with the Universe described by Gaussian normal coordinates.

Conservation of energy-momentum in GR is formulated into the equation

$$T^{ab}{}_{;b} = 0 \qquad \text{or} \qquad \mathfrak{T}^{ab}{}_{;b} = 0 \tag{3.25}$$

where $\mathfrak{T}^{ab} = J\,T^{ab}$ is the energy-momentum density, $J = \sqrt{-det(g_{ab})}$ in general or $J = \sqrt{det(g_{ij})}$ as in section 2.1 for Gaussian normal coordinates. This of course is very different from

$$\mathfrak{T}^{ab}{}_{,b} = 0 \tag{3.26}$$

since the system generally inhabits a non-flat spacetime. If we write (3.26) with $a = 0$, as:

$$\partial_t \int_R \mathfrak{T}^0{}_0\, d^3x = -\int_R \partial_i \mathfrak{T}^i{}_0\, d^3x = \int_{\partial R} \mathfrak{T}^i{}_0\, dS_i \tag{3.27}$$

where integration is evaluated within some 3-region R and its 2-boundary ∂R with dS_i as its surface element; we have the interpretation that the rate of change of a system's total material energy is equal to the total flux of material energy flowing inwards across the boundary. By giving up equation (3.26), do we also lose the conservation of energy? As in the introduction, in FLRW cosmology, we have:

$$dE = -p\, dV \tag{3.28}$$

Indeed the material energy of an arbitrary region therefore the material energy of the whole Universe are decreasing with expansion when the cosmic fluid has non-zero pressure.

The answer to the above puzzle in fact is gravitational energy. GR must reduce to Newtonian gravity in appropriate limit. A system of gravitational bodies in Newtonian gravity do not have conserved kinetic energy in general, as kinetic energy and gravitational potential energy are constantly transferred back and forth between each other. Similarly in GR, we should not expect a conserved material energy in the first place.

Einstein introduced the pseudo-tensor density $\mathfrak{t}^a{}_b$ describing the energy of an arbitrary gravitational field, which has the following expression:

$$\mathfrak{t}^a{}_b = \frac{1}{16\pi}(\delta^a{}_b \mathfrak{L} - \mathfrak{g}^{cd}{}_{,b} \frac{\partial \mathfrak{L}}{\partial \mathfrak{g}^{cd}{}_{,a}}) \tag{3.29}$$

where $\mathfrak{g}^{ab}$ is the metric density and the Lagrangian function is:

$$\mathfrak{L} = J\, g^{ab}(\Gamma^d{}_{ac}\Gamma^c{}_{bd} - \Gamma^c{}_{ab}\Gamma^d{}_{cd}) \tag{3.30}$$

Now we can write (3.25) as

$$\partial_a(\mathfrak{T}^a{}_b + \mathfrak{t}^a{}_b) = 0 \tag{3.31}$$

which of course is just conservation of energy-momentum (3.26) with material energy replaced by the sum of material energy and gravitational energy. However, this proposal, unlike the other aspects of GR, aroused controversy [27]. Those who agreed upon the use of $\mathfrak{t}^a{}_b$ as gravitational energy include Tolman [28], and those who disagreed include Eddinton [29]. The problem lies in the observation that $\mathfrak{t}^a{}_b$ is not a "true" tensor transforming covariantly under coordinate transformation.

As argued by Einstein, the pseudo-tensor $\mathfrak{t}^a{}_b$ has sufficient coordinate-independence in that the total gravitational 4-momentum of an isolated system R calculated as

$$P^a = \int_R (\mathfrak{T}^{0a} + \mathfrak{t}^{0a})\, d^3x \tag{3.32}$$

is independent the choice of coordinate inside the system. By isolated system, we mean a system with no gravitational interaction nor matter interaction with its environment. As pointed out by Einstein, the only meaningful way of talking about an isolated gravitational system is to embed the system in a Minkowski background. This is a good approximation for any system at a scale, say solar scale, such that the cosmological expansion could be neglected. For a finite Universe, it is interesting to ask whether the Universe itself could be regarded as such an isolated system and whether the pseudo-tensor $\mathfrak{t}^{ab}$ would enable us to calculate the total gravitational energy for an exact model. The answer is yes, at least for FLRW model with k=+1. The proof for a general model would be much more difficult, but given the lucid physical argument, it is not unreasonable to postulate the answer is positive.

Still, why is gravitational energy so special? Can we find a better "formula", that is, a tensor field which captures all our physical intuition about gravitational energy? Various effort has been put into solving this problem but no one succeeded. If fact besides Einstein's original expression, there are numerous other suggestions of pseudo-tensorial expression of gravitational energy. Which one should we adopt? They all give the same result as far as integral of energy of an isolated system is concerned. So why there is no tensor field representing gravitational energy? Let us consider three spherical bodies A, B, C in Newtonian gravity, the total gravitational energy is given by

$$E = -\frac{GM_AM_B}{R^2{}_{AB}} - \frac{GM_BM_C}{R^2{}_{BC}} - \frac{GM_CM_A}{R^2{}_{CA}} \tag{3.33}$$

This is different from the sum of the gravitational energy of the system A-B and the gravitational energy of the system B-C. It is completely meaningless to talk about the part of the total gravitational energy belonging to body A or to talk about the part belonging to some spatial volume lying in between A and B. All we can talk about physically is the total gravitational energy of the system. Corresponding to Newtonian gravity, in GR we have the (strong) equivalence principle, which is indeed the spirit and foundation of GR. Given a local volume we can choose a frame where gravity vanishes, therefore "local gravitational field" has no coordinate-free physical meaning thus no tensorial expression [22].

Even the terminology "gravitational field" should be used with much caution. In Newtonian gravity, we can "peel off" a test particle and talk about the gravitational potential at some location outside of a massive gravitational body. However, gravitational potential in GR has no local meaning as argued above. We talk about electromagnetic (EM) field, but the two "fields" are fundamentally different in there nature. EM field can be localized, whereas gravitational field can't. EM serves as part of the right hand side of the Einstein's equation, i.e. material sources bending spacetime; whereas gravitational field does not act as a source bending spacetime since its nature is Geometry[5]. EM field can be quantized

[5]This raises a question: in what sense shall we say gravitational energy has "mass" via the equation $E = mc^2$?

to photon; can gravitational field be quantized? We don't know, but we know a thorough understanding of the non-locality of gravitational field is of crucial importance to the issue.

Before we finish this discussion, we will look at two examples to further illustrate the ideas so far discussed.

In FLRW model, can we interpret (3.28) as "the change of gravitational energy of some co-moving volume element is equal to the work done by the cosmic fluid inside that volume on itself"? Yes and No: we can interpret this as the first law of thermodynamics only applied to the whole Universe but not applied to some arbitrary co-moving region. Why do we have a coordinate-free, conservation of energy type of equation here? This is because of the homogeneity of the cosmological model: equation describing the whole homogeneous Universe must also formally describe any co-moving region, even if a differential region. So does gravitational energy has anything to do with pressure of the fluid? Let us consider irrotational dust in an arbitrary Universe described in Gaussian normal coordinates.

For a co-moving region of dust continuum, equation (3.25) leads directly to (2.14): the energy of the region's dust is conserved. For a co-moving region, we also know that there is no flow of fluid energy across the boundary of the region. Therefore the only possibility is that as long as the pressure of the cosmic fluid is zero there is no transformation of gravitational energy into fluid energy. Intuition may even lead us to postulate that there is no transfer of gravitational energy across the boundary between any two neighbouring co-moving regions; that is: $\partial_t \mathrm{t}^0{}_0 = 0$. Is this true? These arguments only make sense if our co-moving region can be regarded as an isolated system. Can a co-moving region of dust continuum be regarded as an isolated system? After all, idealized dust is a very special kind of object. A straightforward calculation shows the answer is unfortunately no.(See appendix B for details.) We can't regard an arbitrary co-moving region as an isolated system, whether the fluid is simple dust or more general fluid. It makes no sense to to talk about transfer of gravitational energy between two neighbouring co-moving regions.

As commented by Einstein [27]: "The differential law (of conservation of energy) is equivalent to the integral law which has been abstracted from experience; in this alone rests its meaning......We are thus led–contrary to our present habits of thinking—to assign more weight of reality to an integral than to its differentials." We conclude our discussion by emphasizing again that gravitational energy only belongs to the whole isolated system.

Chapter 4

Examples

Applications of the averaging procedure in cosmology developed in the previous two chapters require that the averaging region is compact in order to evaluate appropriate integrals, thus one of the following approaches have to be adopted.

- We can restrict ourselves to the study of a compact portion of the Universe only. But as pointed out in [19], although the averaged equations (2.21),(2.23) and their solutions seem to only depend on the matter distribution and geometry inside the averaging region, the initial data have to be constructed globally. Therefore it is misleading to interpret that an arbitrary patch of the Universe can be described independent of its environment.
- As argued by Wiltshire [30], Buchert's averaging could be applied to the present horizon volume which is compact.
- We can decompose the Universe to an union of compact regions. This method in general depends on an arbitrary choice of decomposition. The study of averaging over different scales and regions in an abstract manner so far has only limited applications, as already pursued in section 2.2 and section 2.3.
- In this chapter we will study another approach used in [19] by Buchert himself, i.e., we **assume** the whole Universe can be modeled as a compact spacelike hypersurface evolving in time. We therefore review the cosmological models provided in [19] as further illustrations to the averaging procedures developed in the previous two chapters, as motivations to new solutions to the system of equations (2.21),(2.23), and we will also look at some of the interesting properties of these models. Through out this chapter, we will drop the subscript of the explicit regional dependence; we emphasize that all spatial averaged quantities in this chapter will be those averaged over the whole compact spacelike hypersurface Σ_t. The assumption of the compactness of the Universe needs to be justified in the first place, but this has not been provided in [19].

4.1 Globally static Universe

Following [19], we will call the Universe with constant scale factor the *globally static Universe*. We have $\dot{a} = \ddot{a} = 0$ and $a = \frac{V(t)}{V_0} = 1$ since the volume is also a constant by the assumption that a is a constant, therefore (2.21),(2.23) now read:

$$\langle R \rangle = 12\pi G \frac{M}{V_0} + 3\Lambda \tag{4.1}$$

$$\mathcal{Q} = 4\pi G \frac{M}{V_0} - \Lambda \tag{4.2}$$

This special solution of course reminds us of Einstein's original static model:

$$4\pi G\rho = \Lambda \qquad \rho = \text{const.} \qquad a = \frac{1}{\sqrt{4\pi G\rho}} \tag{4.3}$$

as a special case of a homogeneous cosmology, with positive spatial curvature. This picture was abandoned not long after it was proposed, firstly due to Hubble's discovery of the redshift-distance relationship which is interpreted as that space defined by the galaxies in our environment is expanding instead of static. As commented in [19], this was actually a hasty decision. The underlying assumption for this abandonment is that the global model of the static Universe could be used to describe any portion of the Universe; the observed portion was regarded as already a fair sample of the entire Universe. In contemporary standard the portion observable back then actually is quite small: in fact, we should also be aware that the current observed Universe could also be only a tiny part of the whole potentially observable Universe.

It was also pointed out by several authors that Einstein's solution is not stable. Following Buchert, we distinguish between global instability and local instability. In the first sense, Einstein's static model is not stable within the class of homogeneous cosmology, since a slight change of density would destroy the balance set up in equation (4.3), leading to a contraction if $4\pi G\rho > \Lambda$, and an expansion otherwise. In the second sense, all homogeneous cosmologies including Einstein's static cosmology are unstable under inhomogeneous density perturbations. Inhomogeneity would be amplified due to the attractive nature of gravity which tends to increase overdensities and to decrease underdensities. In our static solution, however, the difference between these two cases does not matter, since both perturbation will leave the solutions within the same class of cosmologies governed by the averaged equations (2.21) and (2.23). These equations are more general than those of homogeneous cosmology. The perturbation of density does not necessarily destroy the balance (4.1),(4.2). This is because the back-reaction is time dependent in general and is coupled to $\langle R \rangle$, therefore it indeed reacts back to the perturbation. Nevertheless, as concluded in [19], it is still premature to advance the above solutions as a viable model to explain current observations; we mainly study the models as an example of solutions to the averaging problem.

4.2 Globally stationary Universe

In this section, we consider a wider class of solutions than in the last section: the *globally stationary Universe*. The Universe is assumed to have an effective scale factor of the form $a = a_0 + H_0(t - t_0)$, where H_0 is a constant. It is pertinent to mention that the only non-empty stationary homogeneous cosmology is Einstein's static Universe, as is not difficult to see from the Friedmann equation and the Raychaudhuri equation for homogeneous cosmology. We will adopt the normalization for a as in (2.15), i.e., $a_0 = 1$. We also have $\dot{a} = H_0$, $\ddot{a} = 0$ and $H = H_0/a$. Therefore, (2.21) and (2.23) now read:

$$\langle R \rangle = 12\pi G \frac{M}{V_0 a^3} + 3\Lambda - 6H^2 \tag{4.4}$$

$$\mathcal{Q} = 4\pi G \frac{M}{V_0 a^3} - \Lambda \tag{4.5}$$

Inserting (4.5) into (4.4),

$$\langle R \rangle = 3\mathcal{Q} + 6\Lambda - 6H^2 \tag{4.6}$$

and evaluate (4.6) at t_0, we can evaluate the constant H_0 through:

$$6{H_0}^2 = 3\mathcal{Q}_0 + 6\Lambda - \langle R \rangle_0 \tag{4.7}$$

We are now going to derive the exact solution to the averaged equations. First note that the time derivative of (4.6) delivers a dynamical coupling relation between $\mathcal{Q}$ and $\langle R \rangle$ as a special case of the integrability (2.24):

$$-\partial_t \mathcal{Q} + \frac{1}{3}\partial_t \langle R \rangle = \frac{4{H_0}^3}{a^3} \tag{4.8}$$

Evaluate (4.5) at t_0 and then insert it back to (4.5), we solve for $\mathcal{Q}$ to be

$$\mathcal{Q} = \frac{\mathcal{Q}_0 + \Lambda}{a^3} - \Lambda \tag{4.9}$$

We also solve for $\langle R \rangle$ from (4.4), also using (4.6),(4.7) and (4.9), to have:

$$\langle R \rangle = 3\Lambda + \frac{\langle R \rangle_0 - 3\mathcal{Q}_0 - 6\Lambda}{a^2} + \frac{3\mathcal{Q}_0 + 3\Lambda}{a^3} \tag{4.10}$$

We therefore obtain a set of solutions to the averaged equations (2.21) and (2.23) where back-reaction and averaged curvature are dynamically coupled.

We next employ the language of (2.33) to study the globally stationary cosmology. We first define two parameters on which our solutions will depend:

$$\alpha \equiv \Omega_m(t_0) = \frac{2\mathcal{Q}_0 + 2\Lambda}{3{H_0}^2} \qquad \beta \equiv \Omega_\Lambda(t_0) = \frac{\Lambda}{3{H_0}^2} \tag{4.11}$$

where in defining α we have used (4.5). Since from (2.33),

$$\Omega_m \propto \frac{1}{a^3 H^2} \propto \frac{1}{a} \qquad \Omega_\Lambda \propto \frac{1}{H^2} \propto a^2 \tag{4.12}$$

we have:

$$\Omega_m = \frac{\alpha}{a} \qquad \Omega_\Lambda = \beta a^2 \tag{4.13}$$

We write (4.5) and (4.6) in the notation of (2.33) and then use (4.12) to set:

$$\Omega_{\mathcal{Q}} = -\frac{1}{4}\Omega_m + \frac{1}{2}\Omega_\Lambda = -\frac{1}{4}\frac{\alpha}{a} + \frac{1}{2}\beta a^2 \tag{4.14}$$

$$\Omega_{\langle R \rangle} = 1 - \frac{3}{4}\Omega_m - \frac{3}{2}\Omega_\Lambda = 1 - \frac{3}{4}\frac{\alpha}{a} - \frac{3}{2}\beta a^2 \tag{4.15}$$

The cosmic equation of state for the globally stationary cosmology is

$$\frac{p_{tot}}{\rho_{tot}} = \frac{-\frac{\langle R\rangle}{3} + \mathcal{Q}}{\langle R\rangle + \mathcal{Q} - 16\pi G \langle\rho\rangle} = -\frac{1}{3}\Big(\frac{1-3\beta a^2}{1-\beta a^2}\Big) \tag{4.16}$$

which is time dependent for $\Lambda \neq 0$. Although the equation of state approaches $p_{tot} = -\rho_{tot}$ for $|a| \to \infty$, the asymptotic behavior is still stationary instead of a de-Sitter phase, since the cosmological constant only shares the global balance condition (4.7). We also notice that as $|a| \to \infty$, the averaged equations (4.6) and (4.7) will tend to the limit of (4.1),(4.2) in the globally static cosmology

4.3 The example of an accelerating Universe

The globally static and stationary cosmologies were built on the assumption that an exact balance (4.5) is established among the averaged energy density, back-reaction and the cosmological constant. In this section we will consider an example without the restriction of (4.5) and therefore a Universe with acceleration. We will make the simple assumption that the acceleration and the expansion of the Universe are due entirely to the cosmological constant with the matter density and back-reaction satisfying a balance condition as in the case of a static Universe. That is, we decompose the equations (2.21) and (2.23) to four equations:

$$3\,(\frac{\dot a}{a})^2 = \Lambda \qquad -8\pi G\frac{M}{V_0 a^3} + \frac{1}{2}\langle R\rangle - = -\frac{\mathcal{Q}}{2} \tag{4.17}$$

$$3\,\frac{\ddot a}{a} = \Lambda \qquad 4\pi G\frac{M}{V_0 a^3} = \mathcal{Q} \tag{4.18}$$

The equations are easy to solve:

$$a = e^{\sqrt{\frac{\Lambda}{3}}t} \tag{4.19}$$

with $a_0 = 1$. This is just the exponential solution of the flat de-Sitter model. We also have

$$\langle\rho\rangle = \frac{\langle\rho\rangle_0}{a^3} \tag{4.20}$$

Thus

$$\langle R\rangle = 12\pi G\langle\rho\rangle = \frac{\langle R\rangle_0}{a^3} \qquad \mathcal{Q} = 4\pi G\langle\rho\rangle = \frac{\mathcal{Q}_0}{a^3} \tag{4.21}$$

that is, the averaged curvature and back-reaction both obey conservation laws similar to the averaged energy density and the balance equations (4.17) and(4.18) are maintained.

To obtain the density parameters, we first calculate:

$$H = \sqrt{\frac{\Lambda}{3}} \tag{4.22}$$

Using again the parameter $\alpha \equiv \Omega_m(t_0)$, we have:

$$\Omega_m = \frac{\alpha}{a^3} \qquad \Omega_\Lambda = 1 \qquad \Omega_{\langle R\rangle} = -\frac{3}{4}\Omega_m = -\frac{3}{4}\frac{\alpha}{a^3} \qquad \Omega_{\mathcal{Q}} = -\frac{1}{4}\Omega_m = -\frac{1}{4}\frac{\alpha}{a^3} \tag{4.23}$$

The equation of state is the same as that of a globally static Universe with $\Lambda = 0$, i.e.

$$p_{tot} = \rho_{tot} = 0 \tag{4.24}$$

Chapter 5

Conclusion

We re-formulated and examined Buchert's averaging scheme for scalars in GR. The scheme studied here is restricted to a Universe of irrotational dust, using Gaussian normal coordinates; nevertheless a generalisation is possible. Buchert's averaging scheme is mathematically simple. Physically, the averaged quantity as defined in (2.13) has a transparent meaning for the scalar field of energy density ρ via (2.14), and for the scalar field of expansion rate θ via equation (2.18). However, for an arbitrary scalar field, such volume averaging may or may not have a transparent meaning. In particular, consider the following simple analogy: could "wrinkles" on $\mathbb{R}^2$ possibly have the same "average" as "wrinkles" on the surface of 2-sphere S^2? The averaged curvature may indeed serve to characterize the physical intuition of "global curvature", but this needs further justification.

Using the "scalar part" of Einstein's equations as in section 2.1, we re-derived the averaged equations of Buchert. In comparison with the FLRW models, we see that the averaged equations (2.21) and (2.23) are generalisations of the normal Friedmann and Raychaudhuri equations. The averaged equations can be applied to a homogeneous region in which case they reduce back to the familiar form. Indeed, Buchert has put effort into drawing analogy between the generalised averaged equations for an arbitrary region and the Friedmann and Raychaudhuri equations: the averaged curvature $\langle R \rangle$ and the back-reaction of the region $\mathcal{Q}$ effectively can be regarded as extra "source" terms added to Friedmann and Raychaudhuri equations. These analogies seem to be overly formal. In particular, the so called total energy (2.27) is not conserved for a co-moving region and there is no reason to believe it has anything to do with gravitational energy.

By contrast, we cannot apply equations describing a strict homogeneous region in a direct manner to draw overall features of an inhomogeneous region. Equations describing an inhomogeneous region are very different from equations describing a corresponding hypothetical homogeneous region even if the latter has a uniform density equal to the averaged density of the original inhomogeneous region. The averaged equations show the importance of structures in a quantative and clear way; back-reaction plays an important role in cosmology.

The physical reason for the importance of back-reaction may be, as Buchert's averaging shows, that spatial averaging and time evolution do not commute. It was found by Hosoya *et al.* that the source of this non-commutativity is in fact the production rate of relative entropy per unit volume (3.5), and this production rate is bounded by the product of the amplitudes of density and expansion fluctuation (3.7). At first sight, it may seem strange that why should we identify the probability density from the original definition of relative entropy in information theory with our cosmic energy density, but the search for “a deep analogy between thermodynamic and Shannon’s information-theoretic entropy” [31] is indeed a profound open challenge. In this report we have investigated whether the way Hosoya *et al.* cite18 interpret relative entropy can also be applied to Shannon’s entropy. So far this work has very limited use, but the effective entropy defined in section 3.3 does possess certain interesting properties and is worth further study.

Buchert’s averaging requires that the region over which the averaging is performed must be compact, therefore if we do not want to restrict ourselves to a fixed sub-region of the Universe, applications of this averaging procedure in cosmology must either study the same averaging over different scales or, following Buchert himself [19], study cosmology with the assumption that the whole Universe is a compact manifold. Along the first line, we in this project independently found the following interesting property of the information of cosmic inhomogeneity. On further dividing a co-moving region into a family of sub-regions, the total information of the system is the sum of two parts: the first part is the sum of the information of each sub-region(the component information); the second is a quantity describing the inhomogeneity amidst different sub-regions(the structure information). The system is more complicated than the union of its broken components. If we divide the co-moving region further and further, we will have more and more structure information and correspondingly less and less component information.

Along the second line we studied three examples of the application of Buchert’s averaging procedure as illustrations and motivations. We showed that the models constructed based on the averaged equation possess more stability: firstly, amplification of local inhomogeneity due to the attractiveness of gravity would lead to a homogeneous model transforming to inhomogeneous cosmology whereas this is not a problem for an inhomogeneous model governed by more general equations; secondly, a perturbation on the averaged global density would influence the time-dependent back-reaction and the averaged curvature, which may indeed react back to the perturbation, so that any established stability may remain.

We finally point out that Buchert’s averaging scheme have applications in Wiltshire’s cosmological model [32] originally proposed in 2005. The model has recently been improved [30], which focus more closely on the relationship between the average time parameter and proper time of our clocks. in an inhomogeneous cosmology these are not necessarily the same.

We come back to the question asked at the beginning of the introduction, what is the current status of cosmology? Cosmology is a rapidly developing subject. New ideas are proposed constantly. Why has the dark energy become important only during the era when the structure has formed? In his recent paper, Räsänen [33] proposed “the backreaction conjecture”: the relative volume of the regions of space which are expanding faster will come to dominate the slower expanding regions, so that the average expansion rate will rise. Is this valid? We don’t know, but we believe a thorough understanding of the physical meanings of cosmological parameters is vital [32], [30]. Are we close to a cosmology revolution? Again we don’t know, but we know this is an exciting time: is the Universe better described by a matter–curvature–backreaction–dark energy Cosmic Quartet, or better described by a matter–curvature–backreaction Cosmic Trinity?

Acknowledgements

Many thanks to Dr. David. L. Wiltshire, the supervisor of this project, for all his teaching and detailed comments.

I have been lucky enough to receive many kinds of help since I came here in New Zealand. I would especially like to thank the Department of Physics and Astronomy and the Department of Mathematics and Statistics at the University of Canterbury for their generous financial support.

I also would like to use this chance to express my endless love to my parents: my mother HAO JianHua and my father YAO JingBin.

Bibliography

[1] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, (Wiley, 1991).

[2] G. F. R. Ellis, arXiv:astro-ph/0602280.

[3] Spergel *et al.* , arXiv:astro-ph/0603449.

[4] P. J. E. Peebles, Publ. Astron. Soc. Pac. **111** (1999) 274

[5] Y.V.Baryshev, report presented at Gamow Memorial Conference, August 1999, St.-Petersburg, Russia arXiv:astro-ph/9912074.

[6] A. Sandage, G. Tammann and H. Hardy, Ap. J. **172** (1972) 253.

[7] A. Kransiński, *Inhomogeneous Cosmological Models*, (Cambridge, 1997).

[8] L. Pietronero, Physica A. **144** (1987) 257.

[9] M. Joyce *et al.* , Astron. Astrophys. **443** (2005) 11

[10] M. B. Ribeiro, Ap. J. **388** (1992) 1.

[11] A. K. Mittal and D. Lohiya, Fractals **11** (2003) 145.

[12] G. F. R. Ellis and T. Buchert, Phys. Lett. A. **347** (2005) 38.

[13] R. C. Tolman, *Relativity, Thermodynamics and Cosmology*, (Oxford Clarendon, 1934).

[14] R. Zalaletdinov, Annals. Eur. Acad. Sci. (2003) p. 344.

[15] R. Zalaletdinov, Bull. Astron. Soc. India **25** (1997) 401.

[16] G. F. R. Ellis, in B. Bertotti *et al.* (eds.), *General Relativity and Gravitation*, (D. Reidel Publishing, 1984) pp. 215-288.

[17] T. Buchert, Gen. Rel. Grav. **32** (2000) 105.

[18] A. Hosoya, T. Buchert and M. Morita, Phys. Rev. Lett. **92** (2004) 141302.

[19] T. Buchert, Class. Quant. Grav. **23** (2006) 817.

[20] T. Buchert, Gen. Rel. Grav. **33** (2001) 1381.

[21] D. L. Wiltshire, in B. Robson, N. Visvanathan and W. S. Woolcock (eds.), *Cosmology: The Physics of the Universe*, (World Scientific, Singapore, 1996) pp. 473-531.

[22] C. W. Misner, K. S. Thorne and J. A. Wheeler, *Gravitation*, (Freeman, 1973).

[23] R. K. Sachs, in C. DeWitt, B. deWitt(eds.), *Relativity, Groups and Topology*, (Gordon and Breach, 1964) pp. 523-564.

[24] A. Browder, *Mathematical Analysis: an introduction*, (Springer, 1997).

[25] R. Penrose, *The road to reality : a complete guide to the laws of the universe*, (New York : A.A. Knopf, 2005).

[26] M. Amarzguioui and O. Grøn, Phys. Rev. D. **71** (2005) 083011.

[27] A. Einstein, in A. Engel (translation.), *The collected papers of Albert Einstein, volume 7*, (Princeton, 2002), pp. 31-33 and pp. 47-61.

[28] R. C. Tolman, Phys. Rev. Vol. 35, No. 8 (1930) 875.

[29] A. S. Eddinton, *The Mathematical Theory of Relativity, 2nd ed*, (Cambridge: 1954).

[30] D. L. Wiltshire, in preparation.

[31] W. H. Zurek, (eds.), *Complexity, entropy, and the physics of information : the proceedings of the Workshop on Complexity, Entropy, and the Physics of Information held May-June, 1989, in Santa Fe, New Mexico* , (Westview, 1990) p. vii.

[32] D. L. Wiltshire, arXiv:gr-qc/0503099.

[33] S. Räsänen, arXiv:astro-ph/0607626.

[34] L. D. Landau, E. M. Lifshitz, *The Classical Theory of Fields, 4th rev. English. ed, Courses of Theoretical Physics vol 2*, (Pergamon: 1979).

Appendix A

Notation and Conventions

The sign convention in GR is the same as that of [22]. Specifically:

$$\eta_{ab} = (-1, +1, +1, +1) \tag{A.1}$$

$$R^a{}_{bcd} = \Gamma^a{}_{bd,c} - \Gamma^a{}_{bc,d} + \Gamma^a{}_{ec}\Gamma^e{}_{bd} - \Gamma^a{}_{ed}\Gamma^e{}_{bc} \tag{A.2}$$

$$R_{ab} = R^a{}_{acb} \tag{A.3}$$

$$G_{ab} = R_{ab} - \frac{1}{2}g_{ab}R \tag{A.4}$$

Latin letters a, b, c, d (and sometimes e, f, g, h) will be used for 4-dimensional index, where 0 corresponds to timelike coordinate t, and $1, 2, 3$ corresponds to three spatial coordinates. Letters i, j, k (and sometimes l, m, n) will be used for 3-dimensional spatial index.

Semi-colon followed by an index represents covariant differentiation with respect to g_{ab}, e.g., $u_{i;j} \equiv \nabla_j u_i$. "$||$" followed by an index represents covariant differentiation with respect to g_{ij}. Comma followed by an index represents partial differentiation with respect to corresponding coordinate, e.g., $g_{ij,0} \equiv \partial_t(g_{ij})$ and $K_{,i} \equiv \partial_i K$. An over-dot above a scalar represents partial derivative with respect to "t", e.g., $\dot{\theta} \equiv \partial_t \theta$.

When new terminologies appear for the first time, they are labeled in *Italic*.

Appendix B

Details of calculation for section 3.4

We first write down the Christoffel symbols in Gaussian normal coordinates (equation (97.9) of [34]):

$$\Gamma^0{}_{00} = \Gamma^0{}_{i0} = \Gamma^0{}_{0j} = 0, \qquad \Gamma^0{}_{ij} = \frac{1}{2} g_{ij,0} = -K_{ij} \qquad \Gamma^i{}_{0j} = \frac{1}{2} g^{ik} g_{kj,0} = -K^i{}_j \tag{B.1}$$

From (58.52) and the equation following (58.72) of [29], we have

$$\frac{\partial \mathfrak{L}}{\partial \mathfrak{g}^{ab}{}_{,0}} = \delta^0{}_a \Gamma^c{}_{cb} - \Gamma^0{}_{ab} \qquad \mathfrak{g}^{ab}{}_{,0} = J\left(g^{ab}\Gamma^d{}_{0d} - g^{ad}\Gamma^b{}_{d0} - g^{db}\Gamma^a{}_{d0}\right) \tag{B.2}$$

Therefore

$$\begin{aligned}
\frac{\partial \mathfrak{L}}{\partial \mathfrak{g}^{ab}{}_{,0}} \mathfrak{g}^{ab}{}_{,0} &= J\left(g^{00}\Gamma^d{}_{0d}\Gamma^c{}_{c0} - g^{00}\Gamma^b{}_{00}\Gamma^c{}_{cb} - g^{db}\Gamma^0{}_{0d}\Gamma^c{}_{cb}\right. \\
&\qquad \left. - g^{ab}\Gamma^d{}_{0d}\Gamma^0{}_{ab} + g^{ad}\Gamma^b{}_{d0}\Gamma^0{}_{ab} + g^{db}\Gamma^a{}_{d0}\Gamma^0{}_{ab}\right) \\
&= -J\Gamma^i{}_{0i}\Gamma^j{}_{0j} + Jg^{ab}\left(- \Gamma^d{}_{0d}\Gamma^0{}_{ab} + \Gamma^d{}_{b0}\Gamma^0{}_{ad} + \Gamma^d{}_{a0}\Gamma^0{}_{db}\right) \\
&= -J\Gamma^i{}_{0i}\Gamma^j{}_{0j} + Jg^{ij}\left(- \Gamma^k{}_{0k}\Gamma^0{}_{ij} + 2\Gamma^k{}_{i0}\Gamma^0{}_{kj}\right)
\end{aligned} \tag{B.3}$$

On the other hand, (3.30) gives us:

$$\begin{aligned}
\mathfrak{L} &= J\, g^{ij}(\Gamma^b{}_{ia}\Gamma^a{}_{jb} - \Gamma^a{}_{ij}\Gamma^b{}_{ab}) - J\Gamma^j{}_{0i}\Gamma^i{}_{0j} \\
&= J\, g^{ij}(\Gamma^0{}_{ik}\Gamma^k{}_{j0} + \Gamma^k{}_{i0}\Gamma^0{}_{jk} + \Gamma^k{}_{il}\Gamma^l{}_{jk} - \Gamma^0{}_{ij}\Gamma^k{}_{0k} - \Gamma^l{}_{ij}\Gamma^k{}_{lk}) - J\Gamma^j{}_{0i}\Gamma^i{}_{0j} \\
&= J\, g^{ij}(2\Gamma^k{}_{i0}\Gamma^0{}_{jk} - \Gamma^0{}_{ij}\Gamma^k{}_{0k}) + J\, g^{ij}(\Gamma^k{}_{il}\Gamma^l{}_{jk} - \Gamma^l{}_{ij}\Gamma^k{}_{lk}) - J\Gamma^j{}_{0i}\Gamma^i{}_{0j}
\end{aligned} \tag{B.4}$$

Compare (B.3) and (B.4), we see $\partial_t \mathfrak{t}^0{}_0 = 0$ is equivalent to

$$\partial_t(J\, g^{ij}(\Gamma^k{}_{il}\Gamma^l{}_{jk} - \Gamma^l{}_{ij}\Gamma^k{}_{lk})) = \partial_t(J(\Gamma^j{}_{0i}\Gamma^i{}_{0j} - \Gamma^i{}_{0i}\Gamma^j{}_{0j})) \tag{B.5}$$

The Christoffel symbols on the r.h.s. are in fact 3-scalars via of (2.2), whereas the Christoffel symbols on the l.h.s depends on the choice of the coordinates. The only hope for (B.5) to be true is to see whether Einstein's equations for dust can tell us anything about the l.h.s. However the only equation involving $\Gamma^i{}_{jk}$ is equation (2.3), which can also be written as:

$$\Gamma^k{}_{lj}\Gamma^l{}_{0k} - \Gamma^l{}_{0j}\Gamma^k{}_{lk} = \Gamma^i{}_{0j,i} - \Gamma^i{}_{0i,j} \tag{B.6}$$

We now see clearly Einstein's equations, whether for dust or for more general fluid, does not generally lead to vanishing of $\partial_t \mathfrak{t}^0{}_0$.